

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
10 February 2005 (10.02.2005)

PCT

(10) International Publication Number
WO 2005/013131 A2

(51) International Patent Classification⁷: G06F 11/08

(21) International Application Number:

PCT/IB2004/051348

(22) International Filing Date: 30 July 2004 (30.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/492,654 5 August 2003 (05.08.2003) US

(71) Applicant (for all designated States except US): KONINKLIJKE PHILIPS ELECTRONICS, N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(71) Applicant (for AE only): U.S. PHILIPS CORPORATION [US/US]; 1251 Avenue of the Americas, New York, NY 10020 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): ZHAO, Luyin [CN/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(74) Common Representative: KONINKLIJKE PHILIPS ELECTRONICS, N.V.; c/o John Vodopia, P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR PROBABILITY-BASED VALIDATION OF EXTENSIBLE MARKUP LANGUAGE DOCUMENTS

(Schema1)

```
<?xml version="1.0" encoding="utf-8"?>
<xss:schema xmlns:xss="http://www.w3.org/2001/XMLSchema">
  <xss:element name="book">
    <xss:complexType>
      <xss:sequence>
        <xss:element name="title" type="xs:string"/>
        <xss:element name="author" type="xs:string"/>
        <xss:element name="character">
          <xss:complexType>
            <xss:choice>
              <xss:sequence>
                <xss:element name="first-name" type="xs:string"/>
                <xss:element name="friend-of" type="xs:string"/>
                <xss:element name="since" type="xs:date"/>
                <xss:element name="qualification" type="xs:string"/>
              </xss:sequence>
              <xss:sequence>
                <xss:element name="last-name" type="xs:string"/>
                <xss:element name="birth-year" type="xs:string"/>
                <xss:element name="city" type="xs:string"/>
              </xss:sequence>
            </xss:choice>
          </xss:complexType>
        </xss:element>
      </xss:sequence>
      <xss:attribute name="isbn" type="xs:string"/>
    </xss:complexType>
  </xss:element>
</xss:schema>
```

(57) Abstract: A system and method are disclosed to that use a probability-based validation method that looks ahead/back when an incorrect XML tag is found instead of notifying a user about the error immediately. The system and method can provide probability-based values that can be used to point out error locations in a chunk of XML code and indicate most likely error location(s) using probability values.

WO 2005/013131 A2

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:

- *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date
10 February 2005 (10.02.2005)

PCT

(10) International Publication Number
WO 2005/013131 A3

(51) International Patent Classification⁷: G06F 9/45, 17/30

(21) International Application Number: PCT/IB2004/051348

(22) International Filing Date: 30 July 2004 (30.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 60/492,654 5 August 2003 (05.08.2003) US

(71) Applicant (for all designated States except US): KONINKLIJKE PHILIPS ELECTRONICS, N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(71) Applicant (for AE only): U.S. PHILIPS CORPORATION [US/US]; 1251 Avenue of the Americas, New York, NY 10020 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): ZHAO, Luyin [CN/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(74) Common Representative: KONINKLIJKE PHILIPS ELECTRONICS, N.V.; c/o John Vodopia, P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR PROBABILITY-BASED VALIDATION OF EXTENSIBLE MARKUP LANGUAGE DOCUMENTS

(Schema1)

```

<?xml version="1.0" encoding="utf-8"?>
<xss: schema xmlns:xss="http://www.w3.org/2001/XMLSchema">
  <xss: element name="book">
    <xss: complexType>
      <xss: sequence>
        <xss: element name="title" type="xs:string"/>
        <xss: element name="author" type="xs:string"/>
        <xss: element name="character">
          <xss: complexType>
            <xss: choice>
              <xss: sequence>
                <xss: element name="first-name" type="xs:string"/>
                <xss: element name="friend-of" type="xs:string"/>
                <xss: element name="since" type="xs:date"/>
                <xss: element name="qualification" type="xs:string"/>
              </xss: sequence>
              <xss: sequence>
                <xss: element name="last-name" type="xs:string"/>
                <xss: element name="birth-year" type="xs:string"/>
                <xss: element name="city" type="xs:string"/>
              </xss: sequence>
            </xss: choice>
          </xss: complexType>
        </xss: element>
      </xss: sequence>
      <xss: attribute name="isbn" type="xs:string"/>
    </xss: complexType>
  </xss: element>
</xss: schema>

```

(57) Abstract: A system and method are disclosed to that use a probability-based validation method that looks ahead/back when an incorrect XML tag is found instead of notifying a user about the error immediately. The system and method can provide probability-based values that can be used to point out error locations in a chunk of XML code and indicate most likely error location(s) using probability values.



(84) **Designated States (unless otherwise indicated, for every kind of regional protection available):** ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ,*

BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) **Date of publication of the international search report:**

31 March 2005

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.